

# Optimal Filtering of Partially Observed Markov Processes with Gaussian Noise

Liubov A. Markovich

Received: date / Accepted: date

**Abstract** The processing of stationary random sequences under nonparametric uncertainty is given by a filtering problem when the signal distribution is unknown. A useful signal  $(S_n)_{n \geq 1}$  is assumed to be Markovian. This assumption allows us to estimate the unknown  $(S_n)$  using only an observable random sequence  $(X_n)_{n \geq 1}$ . The equation of optimal filtering of such a signal has been received by A.V. Dobrovidov. Our result states that when the unobservable Markov sequence is defined by a linear equation with a Gaussian noise, the equation of optimal filtering coincides both with the classical Kalman filter and the conditional expectation defined by the Theorem on normal correlation.

**Keywords** Markov sequence · The Theorem on normal correlation · Kalman filter · Optimal filtering · Toeplitz matrix

## 1 Introduction

The problem of filtering of unknown signals from the mixture with noise has a wide range of applications including control of linear and nonlinear systems. In the following we consider a partially observable Markov random sequence  $(S_n, X_n)_{n \geq 1}$ , where a useful signal  $S = (S_n)_{n \geq 1}$  is unobservable and the sequence  $X = (X_n)_{n \geq 1}$  is observable. The connection between these variables is given by the following nonlinear (linear) expression

$$X_n = \varphi(S_n, \eta_n), \quad (1)$$

where  $(\eta_n \in \mathbb{R})_{n \geq 1}$  is an i.i.d random sequence,  $(S_n)_{n \geq 1}$  is a Markov sequence and  $\varphi$  is some function. Realizations of random variables (r.v.s)  $S_n \in \mathcal{S}_n \subseteq \mathbb{R}$  and  $X_n \in \mathcal{X}_n \subseteq \mathbb{R}$  are denoted by  $s_1^n = (s_1, \dots, s_n)^T$  and  $x_1^n = (x_1, \dots, x_n)^T$ , respectively.

---

Liubov A. Markovich  
Institute of Control Sciences, Russian Academy of Sciences, Profsoyuznaya 65, 117997 Moscow, Russia  
Tel.: 89168305103  
E-mail: kimo1@mail.ru

In case when (1) has the recursive linear form

$$\begin{aligned} S_n &= aS_{n-1} + b\xi_n, \\ X_n &= AS_n + B\eta_n \end{aligned} \quad (2)$$

where  $S_n, X_n \in \mathbb{R}$  for all  $n$ ;  $\xi_n$  and  $\eta_n$  are mutually independent r.v.s with the standard Gaussian distribution,

$$\begin{aligned} S_0 &\in \mathcal{N}(0, \tilde{\sigma}^2), \quad \tilde{\sigma}^2 = \frac{b^2}{1-a^2} \\ S_n &\in \mathcal{N}(0, 1), \quad n = 1, 2, 3, \dots, \end{aligned}$$

coefficients  $A, B, a, b$  are given real numbers and  $|a| < 1$ , the Kalman filter is applied [Kalman(1960)]. However, the nonlinear models are more important for practice. The extended and unscented Kalman filter algorithms can be applied in this case, [Crisan D.(2013)], [Julier and Uhlmann(2004)].

Another approach for nonlinear processes was proposed in [Stratonovich(1960)]. With this respect, let us define the random sequence  $\vartheta_n = Q(S_n)$ , where the r.v.  $S_n$  is related to  $\vartheta_n \in \Theta_n \subseteq \mathbb{R}$  by some one-to-one function  $Q : \mathcal{S}_n \rightarrow \Theta_n$ . The random sequence  $(\vartheta_n)_{n \geq 1}$  is also a Markov sequence.

To estimate  $\vartheta_n$  the optimal Bayesian estimator in form of the conditional mean

$$\hat{\vartheta}_n = \mathbb{E}(Q(S_n)|x_1^n) = \int_{\mathcal{S}_n} Q(s_n) w_n(s_n|x_1^n) ds_n, \quad (3)$$

has been used. The  $w_n(s_n|x_1^n)$  is the posterior probability density function (pdf). It satisfies the Stratonovich's recurrence equation [Stratonovich(1966)]

$$\begin{aligned} w_1(s_1|x_1) &= \frac{f(x_1|s_1)p(s_1)}{\int_{\mathcal{S}_1} f(x_1|s_1)p(s_1)ds_1}, \\ w_n(s_n|x_1^n) &= \frac{f(x_n|s_n)}{f(x_n|x_1^{n-1})} \int_{\mathcal{S}_{n-1}} p(s_n|s_{n-1}) w_{n-1}(s_{n-1}|x_1^{n-1}) ds_{n-1}, \quad n \geq 2. \end{aligned} \quad (4)$$

Here  $p(s_n|s_{n-1})$  denotes the transition pdf of the Markov sequence  $(S_n)_{n \geq 1}$ ,  $f(x_n|x_1^{n-1})$  and  $f(x_n|s_n)$  denote conditional pdfs.

As the posterior pdf  $w_n(s_n|x_1^n)$  depends on the unknown prior distribution function  $p(s_1)$  and the transition probability  $p(s_n|s_{n-1})$  of the Markov sequence  $(S_n)_{n \geq 1}$ , we cannot use formula (3) to estimate  $\hat{\vartheta}_n$ . To overcome this problem the optimal filtering equation (see Section 2) was proposed in [Dobrovidov(1983)].

We aim to prove the pairwise exact coincidences of the optimal filtering equation in the form (12), Kalman's filter and the conditional expectation  $\mathbb{E}(Q(S_n)|x_1^n)$  defined by the Theorem on normal correlation [Liptser and Shiryaev(2001)]. The latter coincidences are shown to be valid when the unobservable Markov sequence  $(S_n)$  is defined by a linear equation with a Gaussian noise. Thus, the optimal filtering equation is nothing else but the Kalman filter in case of linear model (2). However, for nonlinear processes the optimal filtering equation provides the solution in contrast to the Kalman filter.

The Theorem on normal correlation requires the inverse covariance matrix. In case of the process (2) such matrix has a Toeplitz form [Toeplitz(1907)]. The

explicit inversion of matrices from the Toeplitz class is considered in [Trench(2001), Dow(2003)]. In contrast to classical Toeplitz matrix we deal here with the modified matrix that differs from Toeplitz one by additional term on the diagonal. As an auxiliary result we obtain the explicit inversion of such matrix.

In [Liptser and Shiryaev(2001)] a pseudo inverse matrix was used.

The paper is organized as follows. In Section 2 we remind the general equation of optimal filtering and its special case for the Gaussian pdf. In Section 3 we obtain the conditional density  $f(x_n|x_1^{n-1})$  and its derivative in explicit forms for the linear process (2) and the Gaussian pdf  $f(x_n|s_n)$  (Theorem 1) and show their ratio to estimate  $E(S_n|x_1^n)$ . In Section 4 the coincidence of the general filtration equation for the Gaussian pdf and the Kalman's filter is derived (Theorem 2). In Section 5 we find the explicit inverse covariance matrix  $D_{\mathbf{x}_n, \mathbf{x}_n}^{-1}$  and prove the coincidence of the general filtration equation for the Gaussian pdf and the Theorem on normal correlation (Theorem 3). All proofs are presented in the Appendices.

## 2 Equations of optimal filtering

Motivated by the problem arising in Section 1, we first transform (4) to a form which depends only on known variables.

Integrating (4) over  $s_n$ , we obtain

$$\int_{S_n} w_n(s_n|x_1^n) ds_n = \int_{S_n} \frac{f(x_n|s_n)}{f(x_n|x_1^{n-1})} \int_{S_{n-1}} p(s_n|s_{n-1}) w_{n-1}(s_{n-1}|x_1^{n-1}) ds_{n-1} ds_n.$$

Furthermore, transferring  $f(x_n|x_1^{n-1})$  to the left side of the latter equation we get

$$f(x_n|x_1^{n-1}) = \int_{S_n} f(x_n|s_n) \int_{S_{n-1}} p(s_n|s_{n-1}) w_{n-1}(s_{n-1}|x_1^{n-1}) ds_{n-1} ds_n. \quad (5)$$

Differentiation of (5) in  $x_n$  leads to

$$f'_{x_n}(x_n|x_1^{n-1}) = \int_{S_n} f'_{x_n}(x_n|s_n) \int_{S_{n-1}} p(s_n|s_{n-1}) w_{n-1}(s_{n-1}|x_1^{n-1}) ds_{n-1} ds_n. \quad (6)$$

Let us further assume that the conditional pdf  $f(x_n|s_n)$  belongs to the exponential family of distributions

$$f(x_n|s_n) = \tilde{C}(s_n) h(x_n) \exp(T(x_n)Q(s_n)), \quad (7)$$

where  $\tilde{C}(s_n)$  is a normalization constant and  $h(x_n), T(x_n), Q(s_n)$  are known functions. Its derivative in  $x_n$  is given by

$$f'_{x_n}(x_n|s_n) = f(x_n|s_n) \left( \frac{h'_{x_n}(x_n)}{h(x_n)} + T'_{x_n}(x_n)Q(s_n) \right).$$

Substituting this into (6), we can deduce that

$$\begin{aligned} f'_{x_n}(x_n|x_1^{n-1}) &= \frac{h'_{x_n}(x_n)}{h(x_n)} f(x_n|x_1^{n-1}) \\ &+ T'_{x_n}(x_n) \int_{S_n} f(x_n|s_n) Q(s_n) \int_{S_{n-1}} p(s_n|s_{n-1}) w_{n-1}(s_{n-1}|x_1^{n-1}) ds_{n-1} ds_n. \end{aligned}$$

Dividing the latter equation by  $f(x_n|x_1^{n-1})$  and due to (4) we can write

$$\frac{f'_{x_n}(x_n|x_1^{n-1})}{f(x_n|x_1^{n-1})} = \frac{h'_{x_n}(x_n)}{h(x_n)} + T'_{x_n}(x_n) \int_{S_n} Q(s_n) w_n(s_n|x_1^n) ds_n.$$

Using (3), we can finally write that

$$E(Q(S_n)|x_1^n) \cdot T'_{x_n}(x_n) = \left( \ln \left( \frac{f(x_n|x_1^{n-1})}{h(x_n)} \right) \right)'_{x_n}. \quad (8)$$

This is a general filtration equation obtained in [Dobrovidov(1983)]. Note that equation (8) does not contain the explicit probabilistic characteristics  $p(s_1)$  and  $p(s_n|s_{n-1})$  of the unknown sequence  $(S_n)$ . This allows us to find the optimal estimator (3) knowing only observable quantities of  $x_1^n$ . Further, we shall call (8) as Dobrovidov's equation.

As an example of the exponential family (7) we can take the Gaussian density

$$f(x_n|s_n) = \frac{1}{\sqrt{2\pi}B} \exp \left( -\frac{(x_n - As_n)^2}{2B^2} \right). \quad (9)$$

Then the observation model is defined by the linear equation

$$X_n = AS_n + B\eta_n, \quad (10)$$

where  $\eta_n$  is an i.i.d random sequence with Gaussian distribution and the coefficients  $A$  and  $B$  are given real numbers.

The pdf (9) relates to (7), where

$$\begin{aligned} \tilde{C}(s_n) &= \frac{1}{\sqrt{2\pi}B} \exp \left( -\frac{A^2 s_n^2}{2B^2} \right), \quad h(x_n) = \exp \left( -\frac{x_n^2}{2B^2} \right), \\ T(x_n) &= x_n, \quad Q(s_n) = \frac{s_n A}{B^2}. \end{aligned} \quad (11)$$

Substituting (11) into (8), we can write that

$$E(S_n|x_1^n) = \frac{B^2}{A} \frac{f'_{x_n}(x_n|x_1^{n-1})}{f(x_n|x_1^{n-1})} + \frac{x_n}{A}. \quad (12)$$

The latter formula is a special case of the general filtration equation (8). Furthermore, we need to obtain the conditional density  $f(x_n|x_1^{n-1})$  and its derivative.

### 3 The conditional density $f(x_n|x_1^{n-1})$

In this section we determine the conditional density (5) and its derivative in explicit forms. To this end, we consider a partially observable Markov sequence  $(S_n, X_n)_{n \geq 1}$  defined by the recursive linear equations (2).

The following theorem holds.

**Theorem 1** *The explicit form of the conditional density (5) is defined as*

$$f(x_n|x_1^{n-1}) = \frac{1}{\sqrt{2\pi\sigma_n}} \exp\left(-\frac{1}{2\sigma_n}(x_n - A\mathcal{L}_{n-1})^2\right), \quad n = 2, 3, \dots, \quad (13)$$

where

$$\begin{aligned} \mathcal{L}_n = & \frac{Aa}{\sigma_{n-1}} \left( x_{n-1}\mathfrak{x}_{n-1} + \frac{aB^2}{\sigma_{n-2}} \left( x_{n-2}\mathfrak{x}_{n-2} + \frac{aB^2}{\sigma_{n-3}} \left( x_{n-3}\mathfrak{x}_{n-3} + \dots \right. \right. \right. \\ & \left. \left. \left. + \frac{aB^2}{\sigma_2} \left( x_2\mathfrak{x}_2 + x_1 \frac{aB^2\mathfrak{x}_1}{\sigma_1} \right) \dots \right) \right) \right)^2, \quad n = 2, 3, \dots \end{aligned} \quad (14)$$

with

$$\begin{aligned} \mathfrak{x}_1 &= \tilde{\sigma}^2, \quad \sigma_1 = B^2 + A^2\mathfrak{x}_1, \\ \mathfrak{x}_n &= \frac{B^2 a^2 \mathfrak{x}_{n-1} + \sigma_{n-1} b^2}{\sigma_{n-1}}, \quad \sigma_n = B^2 + A^2\mathfrak{x}_n, \quad n \geq 2. \end{aligned} \quad (15)$$

The proof of Theorem 1 is given in Appendix A.1.

### 3.1 The ratio of the density derivative and the density

Finally, we can find an explicit form of (12). To this end, we have to write the expression for the ratio of the derivative of the density and the density itself. Using Theorem 1, it is straightforward to verify that

$$\begin{aligned} \frac{f'_{x_n}(x_n|x_1^{n-1})}{f(x_n|x_1^{n-1})} = & \frac{1}{\sigma_n} \left( \frac{A^2 a}{\sigma_{n-1}} \left( x_{n-1}\mathfrak{x}_{n-1} + \frac{aB^2}{\sigma_{n-2}} \left( x_{n-2}\mathfrak{x}_{n-2} + \dots \right. \right. \right. \\ & \left. \left. \left. + \frac{aB^2}{\sigma_{n-3}} \left( x_{n-3}\mathfrak{x}_{n-3} + \dots \frac{aB^2}{\sigma_2} \left( x_2\mathfrak{x}_2 + x_1 \frac{aB^2\mathfrak{x}_1}{\sigma_1} \right) \dots \right) \right) \right) - x_n \right) \end{aligned} \quad (16)$$

holds. Substituting (16) into (12), we can write

$$\begin{aligned} \mathbb{E}(S_n|x_1^n) = & \frac{x_n A \mathfrak{x}_n}{\sigma_n} + \frac{x_{n-1} A a B^2 \mathfrak{x}_{n-1}}{\sigma_{n-1} \sigma_n} + \frac{x_{n-2} A a^2 B^4 \mathfrak{x}_{n-2}}{\sigma_{n-2} \sigma_{n-1} \sigma_n} + \dots + \\ & + \frac{x_2 A a^{n-2} B^{2(n-2)} \mathfrak{x}_2}{\sigma_2 \dots \sigma_{n-3} \sigma_{n-2} \sigma_{n-1} \sigma_n} + \frac{x_1 A a^{n-1} B^{2(n-1)} \mathfrak{x}_1}{\sigma_1 \dots \sigma_{n-3} \sigma_{n-2} \sigma_{n-1} \sigma_n}. \end{aligned} \quad (17)$$

Using (14), the ratio (16) can be represented by

$$\frac{f'_{x_n}(x_n|x_1^{n-1})}{f(x_n|x_1^{n-1})} = \frac{A\mathcal{L}_{n-1} - x_n}{\sigma_n}.$$

Then Dobrovodov's equation (17) can be simplified to

$$\mathbb{E}(S_n|x_1^n) = \frac{B^2}{A} \frac{A\mathcal{L}_{n-1} - x_n}{\sigma_n} + \frac{x_n}{A} = \frac{Ax_n \mathfrak{x}_n}{\sigma_n} + \frac{B^2 \mathcal{L}_{n-1}}{\sigma_n}. \quad (18)$$

Considering (18) one can represent (17) in a recursive form. We shall express  $E(S_{n+1}|x_1^{n+1})$  by  $E(S_n|x_1^n)$  using (18). As (14) can be represented as

$$\mathcal{L}_n = \frac{Aa}{\sigma_n} \left( x_n \mathfrak{x}_n + \frac{\mathcal{L}_{n-1} B^2}{A} \right). \quad (19)$$

it can be deduced that

$$\begin{aligned} E(S_{n+1}|x_1^{n+1}) &= \frac{Ax_{n+1}\mathfrak{x}_{n+1}}{\sigma_{n+1}} + \frac{B^2 \mathcal{L}_n}{\sigma_{n+1}} = \\ &= \frac{Ax_{n+1}\mathfrak{x}_{n+1}}{\sigma_{n+1}} + \frac{B^2 \left( \frac{Aa}{\sigma_n} \left( x_n \mathfrak{x}_n + \frac{\mathcal{L}_{n-1} B^2}{A} \right) \right)}{\sigma_{n+1}} \\ &= \frac{Ax_{n+1}\mathfrak{x}_{n+1}}{\sigma_{n+1}} + \frac{B^2 A x_n \mathfrak{x}_n}{\sigma_n \sigma_{n+1}} + \frac{B^2 a}{\sigma_{n+1}} \left( E(S_n|x_1^n) - \frac{Ax_n \mathfrak{x}_n}{\sigma_n} \right). \end{aligned}$$

Therefore, Dobrovydov's equation (12) has a recursive form

$$E(S_{n+1}|x_1^{n+1}) = \frac{Ax_{n+1}\mathfrak{x}_{n+1}}{\sigma_{n+1}} + \frac{B^2 a}{\sigma_{n+1}} E(S_n|x_1^n). \quad (20)$$

Later we shall use this form to prove Theorem 2.

#### 4 The optimal filtering equation and its relation to Kalman filter

Kalman filter for the linear system (2) is defined by following recursive equations [Dobrovidov et al(2012) Dobrovidov, Koshkin, and Vasiliev]

$$E(S_{n+1}|x_1^{n+1}) = \frac{Ab^2 + a^2 A \gamma_n}{B^2 + A^2 b^2 + A^2 a^2 \gamma_n} x_{n+1} + \frac{a B^2 E(S_n|x_1^n)}{B^2 + A^2 b^2 + A^2 a^2 \gamma_n}, \quad (21)$$

$$\gamma_n = \frac{B^2(a^2 \gamma_{n-1} + b^2)}{A^2(a^2 \gamma_{n-1} + b^2) + B^2} \quad (22)$$

under the conditions

$$E(S_1|x_1) = \frac{A\tilde{\sigma}^2}{A^2\tilde{\sigma}^2 + B^2} x_1, \quad \gamma_1 = \frac{B^2\tilde{\sigma}^2}{A^2\tilde{\sigma}^2 + B^2}.$$

The following lemma holds.

**Lemma 1** *The parameters (22) are related to (15) by equation*

$$\gamma_n = \frac{B^2 \mathfrak{x}_n}{\sigma_n},$$

where  $B$  is given by (2).

**Theorem 2** *When a partially observable Markov sequence  $(S_n, X_n)_{n \geq 1}$  is defined by (2), the equation of optimal filtering (12) is equivalent to the Kalman's filter (21).*

Proofs of Lemma 1 and Theorem 2 are given in Appendices A.2 and A.3.

## 5 The Theorem on normal correlation

In [Liptser and Shiryaev(2001)] (Theorem 3.1, p.61) the Theorem on normal correlation has been obtained. For the Gaussian vector  $(\theta, \nu)$  the optimal estimate  $E(\theta|\nu)$  is defined by

$$E(\theta|\nu) = E(\theta) + D_{\theta\nu}D_{\nu\nu}^{-1}(\nu - E(\nu)), \quad (23)$$

where  $E(\theta)$  and  $E(\nu)$  denote expectations and

$$\begin{aligned} D_{\theta\nu} &= \text{cov}(\theta, \nu) = \|\text{cov}(\theta_i, \nu_j)\|, \quad 1 \leq i \leq k, 1 \leq j \leq l \\ D_{\nu\nu} &= \text{cov}(\nu, \nu) = \|\text{cov}(\nu_i, \nu_j)\|, \quad 1 \leq i, j \leq l \end{aligned} \quad (24)$$

are covariance matrices.

The Theorem on normal correlation (23) contains the conditional mathematical expectation as the Dobrovodov's inequality (12). It implies that (12) and (23) can be related. Therefore, we need to find how the covariance matrices (24) can be expressed in terms of (2).

### 5.1 The covariance matrices

From (2) the following conditions

$$\begin{aligned} E(S_0) &= 0, \quad E(S_0^2) = \frac{b^2}{1-a^2}, \\ E(\xi_n) &= 0, \quad E(\eta_n) = 0, \quad E(X_n) = 0, \quad n \geq 1, \\ E(\xi_n^2) &= 1, \quad E(\eta_n^2) = 1, \quad n \geq 1 \end{aligned}$$

follow. Thus, using (2) we can write that  $X_1 = AS_1 + B\eta_1$ , and hence  $S_1 = \frac{X_1 - B\eta_1}{A}$  hold. Then it follows

$$\begin{aligned} S_2 &= aS_1 + b\xi_2 = \frac{a(X_1 - B\eta_1)}{A} + b\xi_2, \\ S_3 &= aS_2 + b\xi_3 = \frac{a^2(X_1 - B\eta_1)}{A} + ab\xi_2 + b\xi_3. \end{aligned}$$

Let  $S_n$  be defined as

$$S_n = \frac{a^{n-1}(X_1 - B\eta_1)}{A} + a^{n-2}b\xi_2 + \dots + a^{n-(n-1)}b\xi_{n-1} + b\xi_n. \quad (25)$$

Then from (2)

$$S_{n+1} = aS_n + b\xi_{n+1} = \frac{a^n(X_1 - B\eta_1)}{A} + a^{n-1}b\xi_2 + \dots + a^{n-n}b\xi_n + b\xi_{n+1}$$

follows. Thus, formula (25) is true for any  $n$  by induction.

Next, we can write down a similar formula for  $X_n$ , i.e. it holds

$$\begin{aligned} X_n &= AS_n + B\eta_n \\ &= a^{n-1}(X_1 - B\eta_1) + a^{n-2}Ab\xi_2 + \dots + aAb\xi_{n-1} + Ab\xi_n + B\eta_n \\ &= Aa^{n-1}(aS_0 + b\xi_1) + a^{n-2}Ab\xi_2 + \dots + aAb\xi_{n-1} + Ab\xi_n + B\eta_n. \end{aligned}$$

Now we turn our attention to covariances.

**Lemma 2** *The following recursive formulas for the covariances*

$$\text{cov}(X_n, X_n) = A^2 \mathfrak{a}_1 + B^2, \quad n \geq 1, \quad (26)$$

$$\text{cov}(X_m, X_n) = A^2 \mathfrak{a}_1 a^{n-m}, \quad n > m, \quad n \geq 1, \quad (27)$$

$$\text{cov}(S_n, X_n) = A \mathfrak{a}_1, \quad \text{cov}(S_n, X_m) = A \mathfrak{a}_1 a^{n-m}, \quad n > m, \quad n \geq 1 \quad (28)$$

hold.

The proof of Lemma 2 is given in Appendix A.4.

Next, combining (26)-(28), we can finally write the covariance matrices of any dimension  $n$

$$\begin{aligned} D_{\mathbf{X}_n, \mathbf{X}_n} &= \begin{pmatrix} \text{var}(X_1) & \text{cov}(X_1, X_2) & \text{cov}(X_1, X_3) & \dots & \text{cov}(X_1, X_n) \\ \text{cov}(X_2, X_1) & \text{var}(X_2) & \text{cov}(X_2, X_3) & \dots & \text{cov}(X_2, X_n) \\ \text{cov}(X_3, X_1) & \text{cov}(X_3, X_2) & \text{var}(X_3) & \dots & \text{cov}(X_3, X_n) \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ \text{cov}(X_n, X_1) & \text{cov}(X_n, X_2) & \text{cov}(X_n, X_3) & \dots & \text{var}(X_n) \end{pmatrix} \\ &= \begin{pmatrix} A^2 \mathfrak{a}_1 + B^2 & A^2 a \mathfrak{a}_1 & A^2 a^2 \mathfrak{a}_1 & \dots & A^2 a^{n-1} \mathfrak{a}_1 \\ A^2 a \mathfrak{a}_1 & A^2 \mathfrak{a}_1 + B^2 & A^2 a \mathfrak{a}_1 & \dots & A^2 a^{n-2} \mathfrak{a}_1 \\ A^2 a^2 \mathfrak{a}_1 & A^2 a \mathfrak{a}_1 & A^2 \mathfrak{a}_1 + B^2 & \dots & A^2 a^{n-3} \mathfrak{a}_1 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ A^2 a^{n-1} \mathfrak{a}_1 & A^2 a^{n-2} \mathfrak{a}_1 & A^2 a^{n-3} \mathfrak{a}_1 & \dots & A^2 \mathfrak{a}_1 + B^2 \end{pmatrix}, \quad (29) \end{aligned}$$

$$\begin{aligned} D_{S_n, \mathbf{X}_n} &= (\text{cov}(S_n, X_1) \text{cov}(S_n, X_2) \dots \text{cov}(S_n, X_n)) \\ &= A \mathfrak{a}_1 (a^{n-1} \ a^{n-2} \ \dots \ 1). \end{aligned} \quad (30)$$

Here  $\mathbf{X}_n = X_1^n = (X_1, \dots, X_n)^T$ .

Matrix  $D_{\mathbf{X}_n, \mathbf{X}_n}$  has to be inverted due to (23). It is not an easy problem to get an explicit matrix inversion. Nevertheless, we further construct the inversion of our covariance matrix (29) for any dimension.

## 5.2 The explicit inversion of Toeplitz matrix $D_{\mathbf{X}_n, \mathbf{X}_n}$

The covariance matrix (29) is called a Toeplitz matrix [Trench(2001)] and can be represented as

$$D_{\mathbf{X}_n, \mathbf{X}_n} = \begin{pmatrix} c_1 + B^2 & ac_1 & a^2 c_1 & \dots & a^{n-1} c_1 \\ ac_1 & c_1 + B^2 & ac_1 & \dots & a^{n-2} c_1 \\ a^2 c_1 & ac_1 & c_1 + B^2 & \dots & a^{n-3} c_1 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ a^{n-1} c_1 & a^{n-2} c_1 & a^{n-3} c_1 & \dots & c_1 + B^2 \end{pmatrix} \quad (31)$$

where  $c_1 = A^2 \mathfrak{a}_1$ . Hence, the covariance matrix (30) can be rewritten as

$$D_{S_n, \mathbf{X}_n} = \frac{c_1}{A} (a^{n-1} \ a^{n-2} \ a^{n-3} \ \dots \ 1). \quad (32)$$



We can represent (31) as follows

$$D_{\mathbf{x}_n, \mathbf{x}_n} = (D_{\mathbf{x}_n, \mathbf{x}_n})_{B=0} + B^2 \mathbf{I}$$

where  $\mathbf{I}$  is the identity matrix and  $(D_{\mathbf{x}_n, \mathbf{x}_n})_{B=0}$  is determined by

$$(D_{\mathbf{x}_n, \mathbf{x}_n})_{B=0} = \begin{pmatrix} c_1 & ac_1 & a^2c_1 & \dots & a^{n-1}c_1 \\ ac_1 & c_1 & ac_1 & \dots & a^{n-2}c_1 \\ a^2c_1 & ac_1 & c_1 & \dots & a^{n-3}c_1 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ a^{n-1}c_1 & a^{n-2}c_1 & a^{n-3}c_1 & \dots & c_1 \end{pmatrix}. \quad (33)$$

Then the inverse matrix reads

$$D_{\mathbf{x}_n, \mathbf{x}_n}^{-1} = ((D_{\mathbf{x}_n, \mathbf{x}_n})_{B=0} + B^2 \mathbf{I})^{-1} = \frac{1}{B^2} \left( \frac{1}{B^2} (D_{\mathbf{x}_n, \mathbf{x}_n})_{B=0} + \mathbf{I} \right)^{-1}.$$

Using the formula  $(\mathbf{P} + \mathbf{I})^{-1} = \mathbf{P}^{-1} - \mathbf{P}^{-1} (\mathbf{I} + \mathbf{P}^{-1})^{-1} \mathbf{P}^{-1}$ , where  $\mathbf{P}$  is a squared invertible matrix [Gantmacher(1990)], we can write that

$$\begin{aligned} D_{\mathbf{x}_n, \mathbf{x}_n}^{-1} &= \frac{1}{B^2} \left( \frac{1}{B^2} (D_{\mathbf{x}_n, \mathbf{x}_n})_{B=0} + \mathbf{I} \right)^{-1} \\ &= (D_{\mathbf{x}_n, \mathbf{x}_n})_{B=0}^{-1} - B^2 (D_{\mathbf{x}_n, \mathbf{x}_n})_{B=0}^{-1} (\mathbf{I} + B^2 (D_{\mathbf{x}_n, \mathbf{x}_n})_{B=0}^{-1})^{-1} (D_{\mathbf{x}_n, \mathbf{x}_n})_{B=0}^{-1}. \end{aligned} \quad (34)$$

To find the inverse matrix  $(D_{\mathbf{x}_n, \mathbf{x}_n})_{B=0}^{-1}$  one can use an algorithm from [Trench(2001)].

Let  $A_n$  be a squared, invertible  $n \times n$  matrix

$$A_n = \begin{pmatrix} 1-a & 0 & 0 & \dots & 0 & 0 \\ 0 & 1 & -a & 0 & \dots & 0 & 0 \\ 0 & 0 & 1 & -a & \dots & 0 & 0 \\ \vdots & \vdots & \vdots & \ddots & \ddots & \vdots & \vdots \\ 0 & 0 & 0 & 0 & \dots & 1 & -a \\ 0 & 0 & 0 & 0 & \dots & 0 & 1 \end{pmatrix}.$$

Then it holds

$$(D_{\mathbf{x}_n, \mathbf{x}_n})_{B=0} \cdot A_n = \begin{pmatrix} c_1 & 0 & 0 & \dots & 0 \\ ac_1 & \alpha_1 & ac_1 & \dots & 0 \\ a^2c_1 & a\alpha_1 & \alpha_1 & \dots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ a^{n-1}c_1 & a^{n-2}\alpha_1 & a^{n-3}\alpha_1 & \dots & \alpha_1 \end{pmatrix}, \quad (35)$$

where  $\alpha_1 = c_1 - a^2c_1$ . Therefore, by multiplying the left and the right sides of (35) by the transposed matrix  $A_n^T$  we can immediately write

$$A_n^T (D_{\mathbf{x}_n, \mathbf{x}_n})_{B=0} A_n = \text{diag}(c_1, \alpha_1, \dots, \alpha_1).$$

Multiplying the latter matrix from the left side by  $(A_n^{-1})^T$  and from the right-hand side by  $A_n^{-1}$ , we obtain

$$(D_{\mathbf{x}_n, \mathbf{x}_n})_{B=0} = (A_n^{-1})^T \text{diag}(c_1, \alpha_1, \dots, \alpha_1) A_n^{-1}.$$

Hence, the inverse matrix is given by

$$(D_{\mathbf{X}_n, \mathbf{X}_n}^{-1})_{B=0} = A_n \text{diag}(c_1^{-1}, \alpha_1^{-1}, \dots, \alpha_1^{-1}) A_n^T.$$

Finally, the inversion of the covariance matrix (33) yields

$$\begin{aligned} (D_{\mathbf{X}_n, \mathbf{X}_n}^{-1})_{B=0} &= \begin{pmatrix} c_1^{-1} + \alpha_1^{-1} a^2 & -\alpha_1^{-1} a & 0 & \dots & 0 \\ -\alpha_1^{-1} a & \alpha_1^{-1} + \alpha_1^{-1} a^2 & -\alpha_1^{-1} a & \dots & 0 \\ 0 & -\alpha_1^{-1} a & \alpha_1^{-1} + \alpha_1^{-1} a^2 & \dots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \dots & \alpha_1^{-1} \end{pmatrix} \\ &= \frac{1}{c_1(1-a^2)} \begin{pmatrix} 1 & -a & 0 & \dots & 0 \\ -a & 1+a^2 & -a & \dots & 0 \\ 0 & -a & 1+a^2 & \dots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \dots & 1 \end{pmatrix}. \end{aligned} \quad (36)$$

Next, using the notation  $d_0 - 1 = \frac{c_1(1-a^2)}{B^2}$  we can write

$$\mathbf{I} + B^2(D_{\mathbf{X}_n, \mathbf{X}_n}^{-1})_{B=0} = \frac{a}{d_0 - 1} \begin{pmatrix} \frac{d_0}{a} & -1 & 0 & \dots & 0 \\ -1 & \frac{d_0+a^2}{a} & -1 & \dots & 0 \\ 0 & -1 & \frac{d_0+a^2}{a} & \dots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \dots & \frac{d_0}{a} \end{pmatrix}.$$

The latter matrix is a tridiagonal, symmetric matrix. In (34) we need its inverse. To this end, we use the theory that was developed in [Fonseca da(2007)], [Usmani(1994)]. Then we have

$$(\mathbf{I} + B^2(D_{\mathbf{X}_n, \mathbf{X}_n}^{-1})_{B=0})^{-1} = \frac{d_0 - 1}{a} \begin{cases} (-1)^{2j} \frac{\psi_{i-1}\varphi_{j+1}}{\psi_n}, & \text{if } i \leq j \\ (-1)^{2i} \frac{\psi_{j-1}\varphi_{i+1}}{\psi_n}, & \text{if } i > j, \end{cases} \quad (37)$$

where  $i, j = 1, \dots, n$  and  $\psi, \varphi$  satisfy the following recurrence relations

$$\psi_m = \left( \frac{d_0 + a^2}{a} \right) \psi_{m-1} - \psi_{m-2}, \quad \text{for } m = 2, \dots, n-1, \quad (38)$$

$$\psi_n = \frac{d_0}{a} \psi_{n-1} - \psi_{n-2}, \quad \text{with initial conditions } \psi_0 = 1, \psi_1 = \frac{d_0}{a}, \quad (39)$$

$$\varphi_k = \left( \frac{d_0 + a^2}{a} \right) \varphi_{k+1} - \varphi_{k+2}, \quad \text{for } k = n-1, \dots, 1$$

$$\text{with initial conditions } \varphi_{n+1} = 1, \varphi_n = \frac{d_0}{a}.$$

Furthermore,  $\psi_m = \varphi_{n+1-m} = \varphi_k, m = 2, \dots, n-1, \quad k = n-1, \dots, 1$ . Thus, (37) can be expressed simply by

$$\begin{aligned} (\mathbf{I} + B^2(D_{\mathbf{x}_n, \mathbf{x}_n})_{B=0}^{-1})^{-1} &= \frac{d_0 - 1}{a\psi_n} \begin{cases} \psi_{i-1}\psi_{n-j}, & \text{if } i \leq j \\ \psi_{j-1}\psi_{n-i}, & \text{if } i > j. \end{cases} \\ &= \frac{d_0 - 1}{a\psi_n} \begin{pmatrix} \psi_{n-1} & \psi_{n-2} & \psi_{n-3} & \dots & \psi_1 & 1 \\ \psi_{n-2} & \psi_1\psi_{n-2} & \psi_1\psi_{n-3} & \dots & \psi_1^2 & \psi_1 \\ \psi_{n-3} & \psi_1\psi_{n-3} & \psi_2\psi_{n-3} & \dots & \psi_2\psi_1 & \psi_2 \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ \psi_1 & \psi_1^2 & \psi_2\psi_1 & \dots & \psi_{n-2}\psi_1 & \psi_{n-2} \\ 1 & \psi_1 & \psi_2 & \dots & \psi_{n-2} & \psi_{n-1} \end{pmatrix}. \end{aligned} \quad (40)$$

Replacing (36) and (40) into (34) one can obtain the explicit inverse covariance matrix (29).

As the product of (32) and (36) is given by

$$D_{S_n, \mathbf{x}_n} \cdot (D_{\mathbf{x}_n, \mathbf{x}_n})_{B=0} = \frac{1}{A} (0 \ 0 \ 0 \ \dots \ 1)$$

then the product of the covariance matrices (32) and (34) is given by

$$\begin{aligned} D_{S_n, \mathbf{x}_n} D_{\mathbf{x}_n, \mathbf{x}_n}^{-1} &= \\ &= -\frac{1}{Aa\psi_n} \begin{pmatrix} 1 - a\psi_1 & -a + (1 + a^2)\psi_1 - a\psi_2 & -a\psi_1 + (1 + a^2)\psi_2 - a\psi_3 & \dots \\ \dots & -a\psi_{n-3} + (1 + a^2)\psi_{n-2} - a\psi_{n-1} & -a\psi_{n-2} + \psi_{n-1} - a\psi_n \end{pmatrix}. \end{aligned}$$

Hence, the Theorem on normal correlation (23) looks as follows

$$\begin{aligned} E(S_n | x_1^n) &= D_{S_n, \mathbf{x}_n} D_{\mathbf{x}_n, \mathbf{x}_n}^{-1} \mathbf{x}_n = \\ &= \frac{a\psi_1 - 1}{Aa\psi_n} x_1 - \frac{a\psi_0 - (1 + a^2)\psi_1 + a\psi_2}{Aa\psi_n} x_2 - \frac{a\psi_1 - (1 + a^2)\psi_2 + a\psi_3}{Aa\psi_n} x_3 - \dots \\ &\dots - \frac{a\psi_{n-3} - (1 + a^2)\psi_{n-2} + a\psi_{n-1}}{Aa\psi_n} x_{n-1} - \frac{a\psi_{n-2} - \psi_{n-1} + a\psi_n}{Aa\psi_n} x_n \end{aligned} \quad (41)$$

## 6 The Theorem on normal correlation and Dobrovidov's equation

Formula (17) can be rewritten as follows

$$\begin{aligned} E(S_n | x_1^n) &= \frac{Aa^{n-1}B^{2(n-1)}\mathfrak{a}_1}{\sigma_1 \dots \sigma_{n-2}\sigma_{n-1}\sigma_n} x_1 + \frac{Aa^{n-2}B^{2(n-2)}\mathfrak{a}_2\sigma_1}{\sigma_1\sigma_2 \dots \sigma_{n-2}\sigma_{n-1}\sigma_n} x_2 + \dots \\ &\dots + \frac{AaB^2\mathfrak{a}_{n-1}\sigma_1\sigma_2 \dots \sigma_{n-2}}{\sigma_1\sigma_2 \dots \sigma_{n-3}\sigma_{n-2}\sigma_{n-1}\sigma_n} x_{n-1} + \frac{A\mathfrak{a}_n\sigma_1\sigma_2 \dots \sigma_{n-2}\sigma_{n-1}}{\sigma_1\sigma_2 \dots \sigma_{n-3}\sigma_{n-2}\sigma_{n-1}\sigma_n} x_n \end{aligned} \quad (42)$$

As we know from (38),  $\psi_n, n = 2, \dots, N-1$  and  $\psi_N$  are described by different formulas. If  $n = N$  is the number of the last element, we would mark the element as the last by  $\tilde{\psi}$ . Then it holds

$$\tilde{\psi}_N = \frac{d_0}{a} \psi_{N-1} - \psi_{N-2}.$$

If the number of the last element is  $n = N + 1$ , then we obtain

$$\psi_N = \frac{d_0 + a^2}{a} \psi_{N-1} - \psi_{N-2} = \tilde{\psi}_N + a \psi_{N-1}.$$

Further, a similar representation can be written for the element  $\psi_{N-1}$

$$\psi_{N-1} = \frac{d_0 + a^2}{a} \psi_{N-2} - \psi_{N-3} = \tilde{\psi}_{N-1} + a \psi_{N-2}$$

and we get

$$\psi_N = \frac{d_0 + a^2}{a} \psi_{N-1} - \psi_{N-2} = \tilde{\psi}_N + a \tilde{\psi}_{N-1} + a^2 \psi_{N-2}.$$

Repeating this procedure we obtain the following formulas

$$\begin{aligned} \psi_N &= \sum_{i=0}^{N-2} \tilde{\psi}_{N-i} a^i + a^{N-1} \psi_1, \\ \psi_{N-1} &= \sum_{i=1}^{N-2} \tilde{\psi}_{N-i} a^{i-1} + a^{N-2} \psi_1 \end{aligned} \quad (43)$$

Then the last element  $\tilde{\psi}_{N+1}$  is the following

$$\begin{aligned} \tilde{\psi}_{N+1} &= \frac{d_0}{a} \psi_N - \psi_{N-1} = \frac{d_0}{a} (\tilde{\psi}_N + a \psi_{N-1}) - \psi_{N-1} = \frac{d_0}{a} \tilde{\psi}_N + (d_0 - 1) \psi_{N-1} \\ &= \frac{d_0}{a} \tilde{\psi}_N + (d_0 - 1) \sum_{i=1}^{N-2} \tilde{\psi}_{N-i} a^{i-1} + (d_0 - 1) a^{N-2} \psi_1 \\ &= \frac{d_0}{a} (\tilde{\psi}_N + (d_0 - 1) a^{N-2}) + (d_0 - 1) \sum_{i=1}^{N-2} \tilde{\psi}_{N-i} a^{i-1}. \end{aligned} \quad (44)$$

The sum in the latter equation is not very convenient. Motivated by this problem we write

$$\begin{aligned} \tilde{\psi}_N &= \frac{d_0}{a} \left( \sum_{i=1}^{N-2} \tilde{\psi}_{N-i} a^{i-1} + a^{N-2} \frac{d_0}{a} \right) - \left( \sum_{i=2}^{N-2} \tilde{\psi}_{N-i} a^{i-2} + a^{N-3} \frac{d_0}{a} \right) \\ &= \sum_{i=1}^{N-2} \tilde{\psi}_{N-i} a^{i-1} \left( \frac{d_0}{a} - \frac{1}{a} \right) + \left( \frac{d_0}{a} \right)^2 a^{N-2} - \frac{d_0}{a} a^{N-3} + \frac{\tilde{\psi}_{N-1}}{a} \end{aligned}$$

where (43) was used. Hence, the sum reads

$$\sum_{i=1}^{N-2} \tilde{\psi}_{N-i} a^{i-1} = \frac{a}{d_0 - 1} \left( \tilde{\psi}_N - \frac{\tilde{\psi}_{N-1}}{a} - \left( \frac{d_0}{a} \right)^2 a^{N-2} + \frac{d_0}{a} a^{N-3} \right) \quad (45)$$

Substituting (45) into (44), we get

$$\begin{aligned} \tilde{\psi}_{N+1} &= \frac{d_0}{a} \tilde{\psi}_N + \frac{d_0}{a} (d_0 - 1) a^{N-2} \\ &\quad + (d_0 - 1) \frac{a}{d_0 - 1} \left( \tilde{\psi}_N - \frac{\tilde{\psi}_{N-1}}{a} - \left( \frac{d_0}{a} \right)^2 a^{N-2} + \frac{d_0}{a} a^{N-3} \right) \\ &= \tilde{\psi}_N \left( \frac{d_0}{a} + a \right) - \tilde{\psi}_{N-1} \end{aligned} \quad (46)$$

**Lemma 3** *If the last element  $\psi_n$  has a number  $n = N$ , where  $N \geq 2$  is an integer number, then*

$$\tilde{\psi}_N = \frac{(1-a^2)}{B^{2N}a^N} \prod_{i=1}^N \sigma_i$$

holds, where  $\tilde{\psi}_N$  is the last element defined by (38).

The proof of Lemma 3 is given in Appendix A.5.

Now we turn our attention to the numerators of (42). Let us introduce the following notations

$$\begin{aligned} C_{x_1} &= Aa^{n-1}B^{2(n-1)}\mathfrak{x}_1 = \frac{b^2}{1-a^2}Aa^{n-1}B^{2(n-1)}, \\ C_{x_i} &= Aa^{n-i}B^{2(n-i)}\mathfrak{x}_i \prod_{j=1}^{i-1} \sigma_j, \quad i = 2, \dots, n. \end{aligned} \quad (47)$$

Parameters (15) can be represented as

$$\mathfrak{x}_n = \frac{B^2a^2\mathfrak{x}_{n-1} + \sigma_{n-1}b^2}{\sigma_{n-1}} = \frac{B^2a^2}{A^2} + b^2 - \frac{B^4a^2}{A^2\sigma_{n-1}}, \quad (48)$$

$$\sigma_n = B^2 + A^2\mathfrak{x}_n = B^2 + A^2b^2 + B^2a^2 - \frac{B^4a^2}{\sigma_{n-1}}, \quad n \geq 2. \quad (49)$$

Hence, we can immediately write

$$\begin{aligned} C_{x_i} &= \frac{aB^2}{A} \left( B^2a^2 + A^2b^2 - \frac{B^4a^2}{\sigma_{i-1}} \right) \prod_{j=1}^{i-1} \sigma_j \\ &= \frac{aB^2}{A} \left( 1 - \frac{B^2}{\sigma_i} \right) \prod_{j=1}^i \sigma_j, \quad i = 2, \dots, n. \end{aligned} \quad (50)$$

Next, the following lemmas can be proved.

**Lemma 4** *The numerators of (41) can be represented as*

$$\begin{aligned} \frac{a\psi_1 - 1}{Aa} &= \frac{Ab^2}{B^2a}, \\ \frac{a\psi_{n-2} - \psi_{n-1} + a\psi_n}{Aa} &= \frac{Ab^2}{B^2a}\psi_{n-1}, \quad n = 2, \dots, N-1, \\ \frac{a\psi_{N-2} - \psi_{N-1} + a\psi_N}{Aa} &= \frac{Ab^2}{B^2a}\psi_{N-1} \end{aligned}$$

The proof of Lemma 4 is given in Appendix A.5.

**Lemma 5** *The numerators of (41) and (42) are related by*

$$\frac{a\psi_1 - 1}{Aa} = \frac{C_{x_1}(1 - a^2)}{B^{2N}a^N}, \quad (51)$$

$$\frac{a\psi_{n-2} - \psi_{n-1} + a\psi_n}{Aa} = \frac{C_{x_n}(1 - a^2)}{B^{2N}a^N}, \quad n = 2, \dots, N-1, \quad (52)$$

$$\frac{a\psi_{N-2} - \psi_{N-1} + a\psi_N}{Aa} = \frac{C_{x_N}(1 - a^2)}{B^{2N}a^N}, \quad (53)$$

where  $C_{x_i}$  is defined by (47).

**Theorem 3** *The theorem on normal correlation (41) and Dobrovidov's equation (42) for the system (2) are coincided.*

The proofs of Lemma 5 and Theorem 3 are given in Appendices A.7 and A.8.

## A Appendix section

### A.1 Proof of Theorem 1

To prove (13) we use mathematical induction. Thus, we have to prove that the statement of Theorem 1 holds for  $n = 2$ . Using (5) we can write

$$f(x_2|x_1) = \int_{S_2} f(x_2|s_2) \int_{S_1} p(s_2|s_1) w_1(s_1|x_1) ds_1 ds_2. \quad (54)$$

The conditional densities  $f(x_1|s_1)$ ,  $f(x_2|s_2)$  defined by (9) are Gaussian. Using the formula (4), where

$$p(s_1) = \frac{1}{\sqrt{2\pi}\tilde{\sigma}} \exp\left(-\frac{s_1^2}{2\tilde{\sigma}^2}\right), \quad \tilde{\sigma}^2 = \frac{b^2}{1 - a^2},$$

we can write the posterior pdf as

$$w_1(s_1|x_1) = \frac{\exp\left(-\frac{(x_1 - As_1)^2}{2B^2} - \frac{s_1^2}{2\tilde{\sigma}^2}\right)}{\int_{S_1} \left(\exp\left(-\frac{(x_1 - As_1)^2}{2B^2} - \frac{s_1^2}{2\tilde{\sigma}^2}\right)\right) ds_1} \quad (55)$$

The integral in the denominator of (55) can be reduced to the form

$$I_{den} = \int_{S_1} \left( \exp\left(-s_1^2 \left(\frac{B^2 + A^2\tilde{\sigma}^2}{2b^2B^2}\right) + s_1 \frac{Ax_1}{B^2} - \frac{x_1^2}{2B^2}\right) \right) ds_1.$$

This is the Euler-Poisson integral that is known in the form

$$\int_{-\infty}^{\infty} \exp(-x^2a^2 + xb + c) dx = \frac{\sqrt{\pi}}{a} \exp\left(\frac{b^2}{4a^2} + c\right). \quad (56)$$

Thus, it is straightforward to verify that

$$I_{den} = \sqrt{\frac{2\pi}{\sigma}} \exp\left(-\frac{x_1^2}{2\sigma B^2 \tilde{\sigma}^2}\right), \quad (57)$$

where  $\sigma = \frac{B^2 + A^2 \tilde{\sigma}^2}{B^2 \tilde{\sigma}^2}$ . Substituting (57) into (55) we deduce the posterior pdf as

$$w_1(s_1|x_1) = \sqrt{\frac{\sigma}{2\pi}} \exp\left(-\frac{\sigma}{2} \left(s_1 - x_1 \frac{A}{B^2 \sigma}\right)^2\right). \quad (58)$$

Since the conditional density in the expression (54) is defined by

$$p(s_2|s_1) = \frac{1}{\sqrt{2\pi}b} \exp\left(-\frac{(s_2 - as_1)^2}{2b^2}\right),$$

we can write using (58) that

$$\int_{s_1} p(s_2|s_1) w_1(s_1|x_1) ds_1 = \frac{1}{\sqrt{2\pi} \mathfrak{a} \mathfrak{e}_2} \exp\left(-\frac{\left(x_1 \frac{Aa\mathfrak{a}_1}{\sigma_1} - s_2\right)^2}{2\mathfrak{a} \mathfrak{e}_2}\right), \quad (59)$$

where the following notations are introduced

$$\mathfrak{a}_1 = \tilde{\sigma}^2, \quad \sigma_1 = B^2 + A^2 \mathfrak{a}_1, \quad \mathfrak{a}_2 = \frac{B^2 a^2 \mathfrak{a}_1 + \sigma_1 b^2}{\sigma_1}, \quad \sigma_2 = B^2 + A^2 \mathfrak{a}_2.$$

Using (59) in (54) and the Euler-Poisson integral (56) we deduce the conditional density for  $n = 2$  as

$$f(x_2|x_1) = \frac{1}{\sqrt{2\pi} \sigma_2} \exp\left(-\frac{1}{2\sigma_2} \left(x_2 - \frac{A^2 a}{\sigma_1} x_1 \mathfrak{a}_1\right)^2\right). \quad (60)$$

Thus, (59) and (60) determine the basis of the mathematical induction.

The second step of the proof is to show that if the following formulas

$$\begin{aligned} \int_{s_{n-1}} p(s_n|s_{n-1}) w_{n-1}(s_{n-1}|x_1^{n-1}) ds_{n-1} &= \frac{1}{\sqrt{2\pi} \mathfrak{a} \mathfrak{e}_n} \exp\left(-\frac{(s_n - \mathcal{L}_{n-1})^2}{2\mathfrak{a} \mathfrak{e}_n}\right), \\ f(x_n|x_1^{n-1}) &= \frac{1}{\sqrt{2\pi} \sigma_n} \exp\left(-\frac{1}{2\sigma_n} \left(x_n - A\mathcal{L}_{n-1}\right)^2\right) \end{aligned} \quad (61)$$

for  $n$  hold, where

$$\begin{aligned} \mathcal{L}_{n-1} &= \frac{Aa}{\sigma_{n-1}} \left( x_{n-1} \mathfrak{a}_{n-1} + \frac{aB^2}{\sigma_{n-2}} \left( x_{n-2} \mathfrak{a}_{n-2} + \right. \right. \\ &\quad \left. \left. + \frac{aB^2}{\sigma_{n-3}} \left( x_{n-3} \mathfrak{a}_{n-3} + \dots \frac{aB^2}{\sigma_2} \left( x_2 \mathfrak{a}_2 + x_1 \frac{aB^2 \mathfrak{a}_1}{\sigma_1} \right) \dots \right) \right) \right), \end{aligned}$$

$\underbrace{\hspace{10em}}_{n-1}$

where  $\mathfrak{a}_n$  and  $\sigma_n$  are defined by (48) and (49), then also formulas (61) are valid for  $n + 1$ .

For  $n + 1$  the posterior density is determined by

$$\begin{aligned} w_n(s_n|x_1^n) &= \frac{f(x_n|s_n)}{f(x_n|x_1^{n-1})} \int_{s_{n-1}} p(s_n|s_{n-1}) w_{n-1}(s_{n-1}|x_1^{n-1}) ds_{n-1} \\ &= \frac{\frac{1}{\sqrt{2\pi}B} \exp\left(-\frac{(x_n - As_n)^2}{2B^2}\right)}{\frac{1}{\sqrt{2\pi} \sigma_n} \exp\left(-\frac{1}{2\sigma_n} \left(x_n - A\mathcal{L}_{n-1}\right)^2\right)} \frac{1}{\sqrt{2\pi} \mathfrak{a} \mathfrak{e}_n} \exp\left(-\frac{1}{2\mathfrak{a} \mathfrak{e}_n} (s_n - \mathcal{L}_{n-1})^2\right) \end{aligned}$$

by its definition. Thus, using (56) and the latter formula we can rewrite (61) for the next step  $n+1$ , i.e.

$$\begin{aligned}
& \int_{S_n} p(s_{n+1}|s_n) w_n(s_n|x_1^n) ds_n = \\
& = \int_{S_n} \frac{\frac{1}{\sqrt{8\pi^3 \mathfrak{a}_n b B}} \exp\left(-\frac{(s_{n+1}-as_n)^2}{2b^2} - \frac{(x_n - A\mathcal{L}_n)^2}{2B^2} - \frac{1}{2\mathfrak{a}_n} \left(s_n - \mathcal{L}_{n-1}\right)^2\right)}{\frac{1}{\sqrt{2\pi\sigma_n}} \exp\left(-\frac{1}{2\sigma_n} \left(x_n - As_{n-1}\right)^2\right)} ds_n ds_{n+1} \\
& = \frac{1}{\sqrt{2\pi\mathfrak{a}_{n+1}}} \exp\left(-\frac{1}{2\mathfrak{a}_{n+1}} \left(s_{n+1} - \left(\frac{Aa}{\sigma_n} x_n \mathfrak{a}_n + \frac{B^2 a \mathcal{L}_{n-1}}{\sigma_n}\right)\right)^2\right) \\
& = \frac{1}{\sqrt{2\pi\mathfrak{a}_{n+1}}} \exp\left(-\frac{1}{2\mathfrak{a}_{n+1}} \left(s_{n+1} - \mathcal{L}_n\right)^2\right)
\end{aligned}$$

holds, where we use the notation (19). Finally, we can write that

$$\begin{aligned}
f(x_{n+1}|x_1^n) &= \int_{S_{n+1}} f(x_{n+1}|s_{n+1}) \int_{S_n} p(s_{n+1}|s_n) w_n(s_n|x_1^n) ds_n ds_{n+1} = \\
&= \int_{S_{n+1}} \frac{\exp\left(-\frac{(x_{n+1}-As_{n+1})^2}{2B^2} - \frac{\left(s_{n+1} - \left(\frac{Aa}{\sigma_n} x_n \mathfrak{a}_n + \frac{B^2 a \mathcal{L}_n}{\sigma_n}\right)\right)^2}{2\mathfrak{a}_{n+1}}\right)}{\sqrt{4\pi^2 \mathfrak{a}_{n+1} B}} ds_{n+1} \\
&= \frac{1}{\sqrt{2\pi\sigma_{n+1}}} \exp\left(-\frac{1}{\sigma_{n+1}} \left(x_{n+1} - A\left(\frac{Aa}{\sigma_n} x_n \mathfrak{a}_n + \frac{B^2 a \mathcal{L}_n}{\sigma_n}\right)\right)^2\right) \\
&= \frac{1}{\sqrt{2\pi\sigma_{n+1}}} \exp\left(-\frac{1}{\sigma_{n+1}} \left(x_{n+1} - A\mathcal{L}_n\right)^2\right).
\end{aligned}$$

Since both the basis and the inductive step have been performed, by mathematical induction, the statement of Theorem 1 holds for all integer  $n > 0$ .

## A.2 Proof of Lemma 1

As the basis we suppose that for  $n=1$  the equation

$$\gamma_1 = \frac{B^2 \tilde{\sigma}^2}{A^2 \tilde{\sigma}^2 + B^2} = \frac{B^2 \mathfrak{a}_1}{\sigma_1}$$

is true.

We have to show as the inductive step that if for  $n$

$$\gamma_n = \frac{B^2 \mathfrak{a}_n}{\sigma_n}$$

holds, then it also holds for  $n+1$ .

By definition we get

$$\gamma_{n+1} = \frac{B^2(a^2\gamma_n + b^2)}{A^2(a^2\gamma_n + b^2) + B^2} = \frac{B^2(a^2\frac{B^2\mathfrak{a}_n}{\sigma_n} + b^2)}{A^2(a^2\frac{B^2\mathfrak{a}_n}{\sigma_n} + b^2) + B^2} = \frac{B^2\mathfrak{a}_{n+1}}{\sigma_{n+1}}$$

Since both the basis and the inductive step have been performed, by mathematical induction the statement of Lemma 1 holds for all integer  $n$ .



## A.3 Proof of Theorem 2

Note, that the denominator of the first term in (21) can be represented as

$$B^2 + A^2 b^2 + A^2 a^2 \gamma_n = B^2 + A^2 \left( \frac{B^2 a^2 \mathfrak{x}_n + \sigma_n b^2}{\sigma_n} \right) = \sigma_{n+1} \quad (62)$$

and its numerator as

$$A b^2 + a^2 A \gamma_n = A \left( \frac{B^2 a^2 \mathfrak{x}_n + \sigma_n b^2}{\sigma_n} \right) = A \mathfrak{x}_{n+1} \quad (63)$$

Thus, (21) can be rewritten using (62), (63) and Lemma 1 as

$$\mathbb{E}(S_{n+1} | x_1^{n+1}) = \frac{A \mathfrak{x}_{n+1}}{\sigma_{n+1}} x_{n+1} + \frac{B^2 a}{\sigma_{n+1}} \mathbb{E}(S_n | x_1^n)$$

that coincides with (20). This implies that Dobrovidov's equation (12) under the condition (2) and the Kalman's filter are coincided.

## A.4 Proof of Lemma 2

We have

$$\begin{aligned} \text{cov}(X_n, X_n) &= \mathbb{E}(X_n - \mathbb{E}(X_n))(X_n - \mathbb{E}(X_n)) = \mathbb{E}(X_n^2) \\ &= \mathbb{E}(A^2 a^{2(n-1)} (a^2 S_0^2 + b^2 \xi_1^2) + a^{2(n-2)} A^2 b^2 \xi_2^2 + \\ &\quad + a^{2(n-3)} A^2 b^2 \xi_3^2 + a^2 A^2 b^2 \xi_{n-1}^2 + A^2 b^2 \xi_n^2 + B^2 \eta_n^2) \\ &= A^2 a^{2(n-1)} \left( a^2 \frac{b^2}{1-a^2} + b^2 \right) + a^{2(n-2)} A^2 b^2 + \\ &\quad + a^{2(n-3)} A^2 b^2 + a^2 A^2 b^2 + A^2 b^2 + B^2 \\ &= A^2 \left( a^{2(n-2)} b^2 \left( \frac{a^2}{1-a^2} + 1 \right) + b^2 a^{2(n-3)} + \dots + a^2 b^2 + b^2 \right) + B^2 \\ &= \frac{A^2 b^2}{1-a^2} + B^2 = A^2 \mathfrak{x}_1 + B^2, \quad n = 1, 2, 3, \dots \end{aligned}$$

Furthermore, it follows

$$\begin{aligned} \text{cov}(X_m, X_n) &= \mathbb{E}(X_m - \mathbb{E}(X_m))(X_n - \mathbb{E}(X_n)) = \mathbb{E}(X_m \cdot X_n) \\ &= A^2 a^{n-1} a^{m-1} (a^2 \mathbb{E}(S_0^2) + b^2 \mathbb{E}(\xi_1^2)) + a^{n-2} a^{m-2} A^2 b^2 \mathbb{E}(\xi_2^2) + \\ &\quad + a^{n-3} a^{m-3} A^2 b^2 \mathbb{E}(\xi_3^2) + a^{m-m} a^{n-m} A^2 b^2 \mathbb{E}(\xi_m^2) \\ &= b^2 A^2 \left( a^{n-2} a^{m-2} \left( \frac{a^2}{1-a^2} + 1 \right) + \dots + a^{n-m} \right) \\ &= \frac{A^2 b^2}{1-a^2} a^{n-m} = A^2 \mathfrak{x}_1 a^{n-m}, \quad n > m, \quad n = 1, 2, 3, \dots \end{aligned}$$

Similarly, we obtain the following two covariances

$$\begin{aligned} \text{cov}(S_n, X_n) &= \frac{A b^2}{1-a^2} = A \mathfrak{x}_1, \\ \text{cov}(S_n, X_m) &= \frac{A b^2}{1-a^2} a^{n-m} = A \mathfrak{x}_1 a^{n-m}, \quad n > m, \quad n = 1, 2, 3, \dots \end{aligned}$$

### A.5 Proof of Lemma 3

Let us assume that for  $\tilde{\psi}_2$  the expression

$$\tilde{\psi}_2 = \frac{(1-a^2)}{B^4 a^2} \sigma_1 \sigma_2$$

is true.

We have to show as the inductive step that if for  $\tilde{\psi}_N$  the equation

$$\tilde{\psi}_N = \frac{(1-a^2)}{B^{2N} a^N} \prod_{i=1}^N \sigma_i, \quad (64)$$

holds than the same holds for  $\tilde{\psi}_{N+1}$ .

Let us substitute (64) into (46). We get

$$\begin{aligned} \tilde{\psi}_{N+1} &= \frac{(1-a^2)}{B^{2N} a^N} \left( \frac{d_0}{a} + a \right) \prod_{i=1}^N \sigma_i - \frac{(1-a^2)}{B^{2(N-1)} a^{N-1}} \prod_{i=1}^{N-1} \sigma_i \\ &= \frac{(1-a^2)}{B^{2(N+1)} a^{N+1}} \prod_{i=1}^{N+1} \sigma_i \left( \frac{B^2 a}{\sigma_{N+1}} \left( \frac{d_0}{a} + a \right) + \frac{B^4 a^2}{\sigma_N \sigma_{N+1}} \right) \\ &= \frac{(1-a^2)}{B^{2(N+1)} a^{N+1}} \prod_{i=1}^{N+1} \sigma_i \left( \frac{\sigma_N \left( B^2 a^2 + A^2 b^2 + B^2 - \frac{B^4 a^2}{\sigma_N} \right)}{\sigma_N \sigma_{N+1}} \right). \end{aligned}$$

Finally, taking into account (49) we can write

$$\tilde{\psi}_{N+1} = \frac{(1-a^2)}{B^{2(N+1)} a^{N+1}} \prod_{i=1}^{N+1} \sigma_i.$$

### A.6 Proof of Lemma 4

By definition we have  $\psi_1 = \frac{d_0}{a} = \frac{A^2 b^2}{B^2 a}$ . Then it follows

$$\frac{a\psi_1 - 1}{Aa} = \frac{Ab^2}{B^2 a}.$$

Using (38) we can immediately write for the second numerator of (41)

$$\begin{aligned} \frac{a\psi_0 - (1+a^2)\psi_1 + a\psi_2}{Aa} &= \frac{a\psi_0 - (1+a^2)\psi_1 + (d_0 + a^2)\psi_1 - a\psi_0}{Aa} \\ &= \frac{-(1+a^2)\psi_1 + \left( \frac{A^2 b^2}{B^2} + 1 + a^2 \right) \psi_1}{Aa} = \frac{Ab^2}{B^2 a} \psi_1. \end{aligned}$$

Similarly, it can be done for any  $n = 2, \dots, N-1$ . For example for  $n = N-1$  we get

$$\begin{aligned} &\frac{a\psi_{N-3} - (1+a^2)\psi_{N-2} + a\psi_{N-1}}{Aa} \\ &= \frac{a\psi_{N-3} - (1+a^2)\psi_{N-2} + (d_0 + a^2)\psi_{N-2} - a\psi_{N-3}}{Aa} \\ &= \frac{-(1+a^2)\psi_{N-2} + \left( \frac{A^2 b^2}{B^2} + 1 + a^2 \right) \psi_{N-2}}{Aa} = \frac{Ab^2}{B^2 a} \psi_{N-2}. \end{aligned} \quad (65)$$

For  $n = N$  the numerator of (41) is different. Using (39) it can be deduced that

$$\begin{aligned} \frac{a\psi_{N-2} - \psi_{N-1} + a\psi_N}{Aa} &= \frac{a\psi_{N-2} - \psi_{N-1} + d_0\psi_{N-1} - a\psi_{N-2}}{Aa} \\ &= \frac{\psi_{N-1}(d_0 - 1)}{Aa} = \frac{Ab^2}{B^2a}\psi_{N-1} \end{aligned}$$

holds.

## A.7 Proof of Lemma 5

For the first numerator of (41) we use (47). Then it is obvious that

$$C_{x_1} = \frac{Ab^2}{B^2a} \frac{a^N B^{2N}}{1 - a^2} = \frac{a\psi_1 - 1}{Aa} \frac{a^N B^{2N}}{1 - a^2} \quad (66)$$

holds. Thus (51) follows. For (52) it is enough to prove this statement for any  $n = \{2, \dots, N-1\}$ . We shall show it for  $n = N-1$ . As we know from Lemma 4 the numerator of (41) for  $n = N-1$  is the following

$$\frac{a\psi_{N-3} - (1 + a^2)\psi_{N-2} + a\psi_{N-1}}{Aa} = \frac{Ab^2}{B^2a}\psi_{N-2}. \quad (67)$$

Therefore, if we use the same technique as in Lemma 3 we can represent  $\psi_{N-2}$  as

$$\begin{aligned} \psi_{N-2} &= \tilde{\psi}_{N-2} + a\psi_{N-3} = \sum_{i=2}^{N-2} \tilde{\psi}_{N-i} a^{i-2} + a^{N-3} \tilde{\psi}_1 \\ &= \sum_{i=2}^{N-2} \tilde{\psi}_{N-i} a^{i-2} + a^{N-4} d_0 \end{aligned} \quad (68)$$

Using the similar technique as in (46) we get

$$\begin{aligned} \tilde{\psi}_{N-1} &= \frac{d_0}{a} \left( \sum_{i=2}^{N-2} \tilde{\psi}_{N-i} a^{i-2} + a^{N-3} \frac{d_0}{a} \right) - \left( \sum_{i=3}^{N-2} \tilde{\psi}_{N-i} a^{i-3} + a^{N-4} \frac{d_0}{a} \right) \\ &= \sum_{i=2}^{N-2} \tilde{\psi}_{N-i} a^{i-2} \left( \frac{d_0}{a} - \frac{1}{a} \right) + \left( \frac{d_0}{a} \right)^2 a^{N-3} - \frac{d_0}{a} a^{N-4} + \frac{\tilde{\psi}_{N-2}}{a}. \end{aligned}$$

Next, expressing from the latter equation the sum

$$\sum_{i=2}^{N-2} \tilde{\psi}_{N-i} a^{i-2} = \frac{a}{d_0 - 1} \left( \tilde{\psi}_{N-1} - \frac{\tilde{\psi}_{N-2}}{a} - \left( \frac{d_0}{a} \right)^2 a^{N-3} + \frac{d_0}{a} a^{N-4} \right)$$

and substituting it into (68) we can write

$$\begin{aligned} \psi_{N-2} &= \frac{a}{d_0 - 1} \left( \tilde{\psi}_{N-1} - \frac{\tilde{\psi}_{N-2}}{a} - \left( \frac{d_0}{a} \right)^2 a^{N-3} + \frac{d_0}{a} a^{N-4} \right) + d_0 a^{N-4} \\ &= \frac{a\tilde{\psi}_{N-1} - \tilde{\psi}_{N-2}}{d_0 - 1} = \frac{B^2(a\tilde{\psi}_{N-1} - \tilde{\psi}_{N-2})}{A^2 b^2} \\ &= \frac{B^2(1 - a^2)}{A^2 b^2} \left( \frac{a \prod_{j=1}^{N-1} \sigma_j}{B^{2(N-1)} a^{N-1}} - \frac{\prod_{j=1}^{N-2} \sigma_j}{B^{2(N-2)} a^{N-2}} \right) \end{aligned}$$

We substitute the latter results into (67). Finally, we deduce that

$$\begin{aligned} \frac{a\psi_{N-3} - (1+a^2)\psi_{N-2} + a\psi_{N-1}}{Aa} &= \frac{(1-a^2)}{AB^{2(N-1)}a^{N-1}} \left(1 - \frac{B^2}{\sigma_{N-1}}\right) \prod_{j=1}^{N-1} \sigma_j \\ &= C_{x_{N-1}} \frac{(1-a^2)}{B^{2N}a^N} \end{aligned}$$

holds, where the definition (50) of  $C_{x_{N-1}}$  is used. Thus, the statement (52) is proved. For the last case when  $n = N$  the same results are valid. From Lemma 4 we get that

$$\frac{a\psi_{N-2} - \psi_{N-1} + a\psi_N}{Aa} = \frac{Ab^2}{B^2a} \psi_{N-1} \quad (69)$$

holds. From (43), (45) it follows

$$\psi_{N-1} = \frac{a\tilde{\psi}_N - \tilde{\psi}_{N-1}}{d_0 - 1} = \frac{B^2(1-a^2)}{A^2b^2} \left( \frac{a \prod_{j=1}^N \sigma_j}{B^{2N}a^N} - \frac{\prod_{j=1}^{N-1} \sigma_j}{B^{2(N-1)}a^{N-1}} \right)$$

Substituting it into (69) and using (50) we deduce

$$\frac{a\psi_{N-2} - \psi_{N-1} + a\psi_N}{Aa} = \frac{(1-a^2)}{AB^{2N}a^N} \left(1 - \frac{B^2}{\sigma_N}\right) \prod_{j=1}^N \sigma_j = C_{x_N} \frac{(1-a^2)}{B^{2N}a^N}.$$

## A.8 Proof of Theorem 3

The assertion of the theorem follows immediately from Lemmas 3 and 4.

**Acknowledgements** L. A. M. acknowledges the financial support provided within the Russian Foundation for Basic Research, grant 13-08-00744. A. L. A. M. would like to thank Prof. A.V. Dobrovidov and Prof. R.S. Liptser for reading a draft of this paper and giving useful comments.

## References

- Crisan D.(2013). Crisan D MJ (2013) Particle approximation of the filtering density for state-space markov models in discrete time. <http://arxiv.org/abs/11115866>
- Dobrovidov(1983). Dobrovidov AV (1983) Nonparametric methods of nonlinear filtering of stationary random sequences. *Automat and Remote Control* 44 (6):757 – 768
- Dobrovidov et al(2012)Dobrovidov, Koshkin, and Vasiliev. Dobrovidov AV, Koshkin GM, Vasiliev VA (2012) Non-Parametric state space models. Kendrick press, USA
- Dow(2003). Dow M (2003) Explicit inverses of toeplitz and associated matrices. *ANZIAM J* 44 (E):E185–E215
- Fonseca da(2007). Fonseca da CM (2007) On the eigenvalues of some tridiagonal matrices. *J Comput Appl Math* 200:283 – 286
- Gantmacher(1990). Gantmacher FR (1990) *The Theory of Matrices*
- Julier and Uhlmann(2004). Julier SJ, Uhlmann JK (2004) Unscented filtering and nonlinear estimation. *Proceedings of the IEEE* 92 (3):401–422
- Kalman(1960). Kalman RE (1960) A new approach to linear filtering and prediction problems. *Journal of Basic Engineering* 82(1):35 – 45
- Liptser and Shiryaev(2001). Liptser RS, Shiryaev AN (2001) *Statistics of Random Processes: II. Applications*. Springer

- Stratonovich(1960). Stratonovich R (1960) Conditional markov processes. *Theory of Probability and its Applications* 5:156 – 178
- Stratonovich(1966). Stratonovich RL (1966) Conditional Markovian processes and their application to the optimal control theory. Moscow Univ. Press, (in Russian).
- Toeplitz(1907). Toeplitz O (1907) Zur transformation der scharen bilinearer formen von unendlichvielen veränderlichen. *Nachrichten der Kgl Gesellschaft der Wissenschaften zu Göttingen, mathematisch-physikalische Klasse* pp 110–115
- Trench(2001). Trench WF (2001) Properties of some generalizations of Kac-Murdock-Szegö matrices, *American Mathematical Society, Providence, RI*, pp 233–246. *Contemporary mathematics*
- Usmani(1994). Usmani RA (1994) Inversion of jacobi's tridiagonal matrix. *Computers Math Applic* 27(8):59 – 66

